

Attachment 12

Experimental Study of Cigarette Warnings: Study 2 Report

OMB Control No. 0910-0866

May 2019

Table of Contents

Section	Page
Executive Summary of Methods and Results	2
Methodology Report	7
Results Report	81
Supplemental Analyses	159

This information is distributed solely for the purpose of the pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by FDA. It does not represent and should not be construed to represent any agency determination or policy.

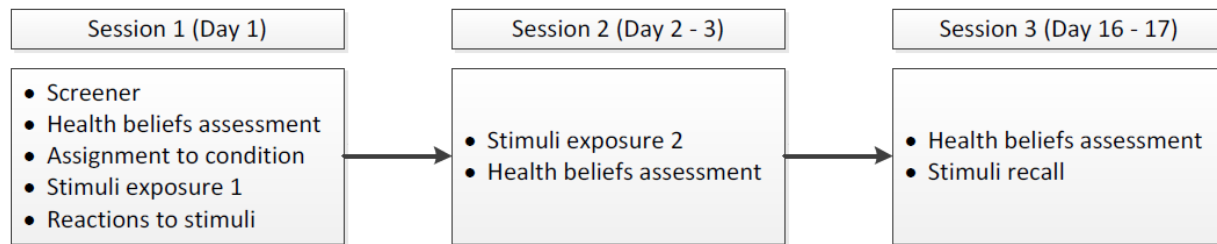
**Executive Summary of Methods and Results for
*Experimental Study Experimental Study of Cigarette Warnings (OMB# 0910-0866)***

Background: Under Section 201 of the Family Smoking Prevention and Tobacco Control Act (TCA) (Pub. L. 111-31), FDA is required to issue a regulation requiring color graphics that depict the negative health consequences of cigarette smoking to accompany new textual warning statements required by the TCA. Pursuant to Section 202(b) of the TCA, the Secretary may adjust the text of the TCA warning statements if doing so would “promote greater public understanding of the risks associated with the use of tobacco products.” As part of the new cigarette health warning development process, FDA developed revised textual warning statements that were tested in a previous study (OMB# 0910-0848) along with the warning statements provided in the TCA. Based on the results of that study, FDA selected 15 warning statements that were then paired with concordant photorealistic images that depicted the negative health consequences of cigarette smoking to form 16 cigarette health warnings to be tested in the present study. Those images were developed and tested through previous formative research.

Purpose of the Study: The main goal of this study was to assess which, if any, of the cigarette health warnings being tested promote greater understanding of the risks associated with cigarette smoking as compared to the Surgeon General’s warnings (i.e., the *status quo* currently on cigarette packages and in cigarette advertisements).

Participants Included: This study included 9,760 participants recruited through an existing online panel called Lightspeed. There were 2,301 adolescents (ages 13-17 years), including both current smokers and those who had never smoked but were at risk for starting smoking. There were 2,701 young adult (ages 18-24 years) current smokers and nonsmokers and 5,388 older adult (ages 25 years and older) current smokers and nonsmokers.

Design of the Study: Participants in all age groups were randomly assigned to a condition that determined which warning they viewed during the study. Participants in the control condition viewed 1 of the 4 current Surgeon General’s warnings; participants in each of the 16 treatment conditions viewed 1 of 16 new cigarette health warnings. All participants viewed their assigned warning on both a mock cigarette pack and a mock cigarette magazine advertisement.

Study Procedure:

The study had three sessions. In Session 1, participants completed a baseline assessment of health beliefs, then viewed their assigned warnings on both the mock cigarette pack and mock cigarette advertisement and then completed a set of questions assessing their reactions to the warnings. In Session 2, one to two days after Session 1, participants viewed the same warnings they saw during Session 1 and completed a second set of questions assessing health beliefs. In Session 3, approximately 14 days after Session 2, participants responded to a final set of questions assessing health beliefs and an item measuring recall of the warnings they viewed in the previous sessions.

Questions were designed to measure several study outcomes, including:

- whether the warning was new information to participants (*"New information"*);
- whether participants learned something from the warning (*"Self-reported learning"*);
- whether the warning made participants think about the health risks of smoking (*"Thinking about risks"*);
- whether the warning was perceived to be informative (*"Perceived Informativeness"*);
- whether the warning was perceived to be understandable (*"Perceived Understandability"*);
- whether the warning was perceived to be a fact or opinion (*"Factualness"*);
- beliefs about the link between smoking and each of the health consequences presented in the warning (*"Health beliefs"*);
- whether the warning was perceived to help participants understand the negative health effects of smoking (*"Perceived helpfulness understanding health effects"*);
- whether the warning grabbed their attention (*"Attention"*);
- whether the warning was recalled (*"Recall"*).

Overview of Statistical Analyses: Analyses compared the responses from participants in each of the treatment conditions to responses from participants in the control condition (i.e., average of the 4 Surgeon General's warnings). These analyses examined whether, relative to viewing a Surgeon General's warning, viewing a cigarette health warning resulted in statistically significantly higher levels of the outcomes measured (e.g., *New information, Self-reported learning*). For the change in health beliefs over the sessions these analyses examined whether, relative to changes in level of agreement with the health belief statements between session for those participants in the control condition, the difference between sessions was larger for those participants in each of the treatment conditions (cigarette health warnings).

Summary of Results: In general, the vast majority of the new cigarette health warnings tested showed statistically significant effects across the outcomes measured, as compared to the current Surgeon General's warnings. Participants were significantly more likely, relative to the control condition (i.e., the Surgeon General's warnings), to report that for 13 of the 16 cigarette health warnings tested (except for Addictive, Kill you, and Quit now): the new cigarette health warnings provided *new information*, were higher on *self-reported learning*, and that the new cigarette health warnings were higher on *perceived informativeness*. Participants in nearly all cigarette health warning conditions (15 of 16) were significantly more likely, relative to the control condition, to rate the warnings as higher on *Perceived Understandability* (except for Quit now) and to report that the warnings were higher on *Perceived helpfulness understanding health effects* (except for Addictive). Similarly, participants in 14 of the 16 cigarette health warning conditions rated these warnings statistically significantly higher on *thinking about risks* (except for Addictive and Quit now), relative to the control condition. All warnings (new cigarette health warnings and current Surgeon General's warnings) were rated as factual by the vast majority of participants, however only half of the cigarette health warnings (8 of 16) were rated as higher on *Factualness* by more participants relative to the control condition. Participants in all 16 cigarette health warnings conditions were more likely, relative to participants in the control condition, to report that the warning they viewed would attract *attention*. Participants in all 16 cigarette health warnings conditions were more likely to be accurately *recall* which warning they had seen than were participants in the control condition. As for changes in *health beliefs*, between Session 1 and Session 2 (approximately 1-2 days apart), 11 cigarette health warnings resulted in greater net positive changes in participants' agreement with health belief items linking smoking to a specific health consequence, and

2. Study Design

2.1 Experimental Design

Participants from six subgroups (adolescent smokers, adolescents susceptible to smoking, young adult smokers, young adult nonsmokers, older adult smokers, and older adult nonsmokers) were randomized to 1 of 16 treatment conditions viewing GHW or a control condition viewing Surgeon General's (SG) warnings. Specifically, within each of the six subgroups, the participants were assigned to the condition with the lowest count (i.e., fewest number of participants), and if more than one condition shared the lowest count, the participants were randomly assigned to one of those conditions. Participants remained in the same condition throughout the study. We also set recruitment and sampling parameters to minimize skewing in demographic characteristics that we anticipated could differ substantially in the sample as compared with population distributions. Specifically, we had quotas such that female participants would comprise no more than 60% of the total sample, and adults aged 65 and older would comprise no more than 25% of the total older adult (aged 25+) sample.

The experimental conditions appear in Table 2-1. In the control condition (condition 0), participants saw a random selection of one of four SG warnings displayed on an image of a mock cigarette pack and mock cigarette advertisement (order of pack and advertisement was randomized). In the treatment conditions (conditions 1-16), participants saw a GHW (i.e., a warning statement combined with an image) displayed on an image of a mock cigarette pack and mock cigarette advertisement (order of pack and advertisement was randomized). Conditions 10 and 11 used the same statement ("Smoking causes COPD, a lung disease that can be fatal.") but with different images (diseased lungs or man with oxygen); because the images differed, conditions 10 and 11 were treated as distinct warnings. Appendix A displays the stimuli for all conditions.

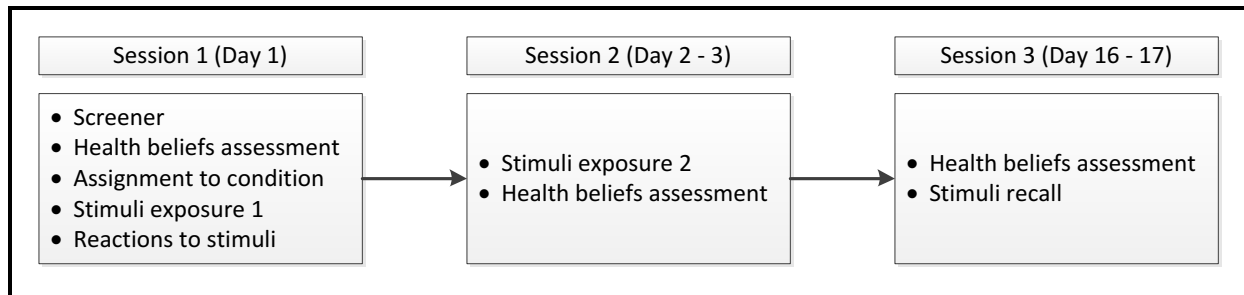
Table 2-1. Study Conditions

Condition Number	Exposure^a	Abbreviated Term for Warning
Surgeon General's Warnings		
0 (control)	Random selection of 1 of the following SG warnings:	
	1) SURGEON GENERAL'S WARNING: Smoking Causes Lung Cancer, Heart Disease, Emphysema, and May Complicate Pregnancy.	SG Disease
	2) SURGEON GENERAL'S WARNING: Quitting Smoking Now Greatly Reduces Serious Risks to Your Health.	SG Quitting
	3) SURGEON GENERAL'S WARNING: Smoking by Pregnant Women May Result in Fetal Injury, Premature Birth, and Low Birth Weight.	SG Pregnancy
	4) SURGEON GENERAL'S WARNING: Cigarette Smoke Contains Carbon Monoxide.	SG Carbon Monoxide
Graphic Health Warnings		
1	WARNING: Cigarettes are addictive.	Addictive
2	WARNING: Tobacco smoke can harm your children.	Harm children
3	WARNING: Smoking can kill you.	Kill you
4	WARNING: Tobacco smoke causes fatal lung disease in nonsmokers.	Fatal lung disease in nonsmokers
5	WARNING: Quitting smoking now greatly reduces serious risks to your health.	Quit now
6	WARNING: Smoking causes head and neck cancer.	Head and neck cancer
7	WARNING: Smoking causes bladder cancer, which can lead to bloody urine.	Bladder cancer
8	WARNING: Smoking during pregnancy stunts fetal growth.	Stunt fetal growth
9	WARNING: Smoking can cause heart disease and strokes by clogging arteries.	Clogged arteries
10	WARNING: Smoking causes COPD, a lung disease that can be fatal. [Paired with COPD Image 1: diseased lungs]	COPD 1
11	WARNING: Smoking causes COPD, a lung disease that can be fatal. [Paired with COPD Image 2: man with oxygen]	COPD 2
12	WARNING: Smoking reduces blood flow, which can cause erectile dysfunction.	Erectile dysfunction
13	WARNING: Smoking reduces blood flow to the limbs, which can require amputation.	Amputation
14	WARNING: Smoking causes type 2 diabetes, which raises blood sugar.	Diabetes
15	WARNING: Smoking causes age-related macular degeneration, which can lead to blindness.	Macular degeneration
16	WARNING: Smoking causes cataracts, which can lead to blindness.	Cataracts

^a Stimuli (warnings on mock cigarette packs and ads) appear in Appendix A. SG = Surgeon General's Warning.

The study comprised three Sessions, outlined in Figure 2-1.

Figure 2-1. Study Protocol



In Session 1, which took approximately 12 minutes, participants first completed a screening questionnaire through an email invitation. After screening for inclusion (see Study Screener in Appendix A of the Final Results Report), participants who qualified for the study completed three consecutive components: (1) a baseline assessment of beliefs about the negative health consequences of cigarette smoking (i.e., “health beliefs”); (2) assignment to study condition and exposure to cigarette warning stimuli according to condition assignment; and (3) assessment of new information, self-reported learning, and other reactions to the stimuli (see Session 1 Survey Instrument in Appendix A of the Final Results Report). These three components are described below.

- **Component (1):** First, participants were asked questions about beliefs related to the health consequences of cigarette smoking.
- **Component (2):** Following the baseline assessment of health beliefs, participants were randomized to one of 16 treatment conditions or a control condition with variation in exposure to cigarette warnings. Participants in each treatment condition were exposed to one GHW, with each condition corresponding to a unique warning from a set of 16. Participants in the control condition were exposed to a random selection of one of four SG warnings. Each stimuli exposure included viewing of the warning in two formats: on a mock cigarette package depicted in a 3-dimensional, rotational model; and on a mock cigarette advertisement. The order of viewing the package and advertisement formats were randomized. In all analyses, stimuli exposure was considered the joint exposure to both stimuli formats; stimuli format was not considered a study factor.
- **Component (3):** After viewing the warning stimuli in both package and advertisement formats, participants completed a brief set of measures to assess (a) if the information presented in the warning was new; (b) self-reported learning from the warning; (c) understandability of the warning; (d) if the warning was perceived to be a fact or an opinion; (e) informativeness of the warning; (f) if the warning grabbed their attention; and (g) if the warning made them think about the health risks of smoking.

One to two days following completion of the baseline assessment (Session 1), Session 1 participants received an email invitation to complete a follow-up (Session 2). In this follow-up session (approximately 8 minutes in duration), participants were re-exposed to the warning stimuli they were shown in Session 1. This exposure followed the same protocol described in Component 2, above. Following stimuli exposure, participants completed a set of immediate post-test measures assessing beliefs related to the negative health consequences of cigarette smoking (see Session 2 Survey Instrument in Appendix B of the Final Results Report).

Fourteen days after Session 2, at the delayed post-test (Session 3, approximately 5 minutes in duration), Session 2 participants received an email invitation to complete a questionnaire assessing measures of beliefs about the negative health consequences of cigarette smoking, as well as recall of the warning (see Session 3 Survey Instrument in Appendix C of the Final Results Report).

2.2 Sampling Frame and Sampling Methodology

Study participants were recruited from a national online panel of adults managed by Lightspeed. The Lightspeed panel is a non-probability convenience sample recruited via social media, online recruitment (e.g., via banner placements), and affiliate corporate networks. For the current study, Lightspeed recruited adult panelists and parents of potential youth respondents using information from panelists' user profiles related to study eligibility (i.e., age, smoking status, and whether or not the panelist had a child in the eligible age range). Recruitment focused on six subgroups (adolescent smokers, adolescents susceptible to smoking, young adult smokers, young adult nonsmokers, older adult smokers, and older adult nonsmokers) based on the criteria listed in Table 2-2.

Table 2-2. Age and Smoking-Related Criteria for Inclusion in Subgroup

Subgroup	Age	Smoking-Related Criteria
Adolescent smokers	13–17	Smoked a cigarette in past 30 days
Adolescents susceptible to smoking	13–17	Never tried cigarettes and responded anything <u>other</u> than “definitely not” to ≥1 of 4 questions assessing susceptibility. (Pierce et al., 1996)
Young adult smokers	18–24	Smoked 100 cigarettes in lifetime and now smoke “every day” or “some days”
Young adult nonsmokers	18–24	Now smoke “not at all”
Older adult smokers	≥25	Smoked 100 cigarettes in lifetime and now smoke “every day” or “some days”
Older adult nonsmokers	≥25	Now smoke “not at all”

3.2 Phase 1 Results

3.2.1 New Information

As shown in Table 3-3, 27.9% of participants in the control condition described the warning as new information; between 22.8% (Addictive) and 88.7% (Cataracts) participants in treatment conditions described the warning they viewed as new information. Participants were significantly more likely to describe a warning as providing new information relative to the control in 13 conditions: Harm children, Fatal lung disease in nonsmokers, Head and neck cancer, Bladder cancer, Stunt fetal growth, Clogged arteries, COPD 1, COPD 2, Erectile dysfunction, Amputation, Diabetes, Macular degeneration, and Cataracts. One warning (Addictive) was less likely to be considered new information than the control condition. These comparisons were statistically significant before and after controlling for multiple comparisons.

Table 3-3. Logistic Regressions of New Information, Thinking about Risks, and Perceived Factualness Comparing GHW with SG Warnings

Condition	Warning	New Information		Thinking About Risks		Perceived Factualness	
		%	OR (95% CI)	%	OR (95% CI)	%	OR (95% CI)
0 (Control)	Average of the 4 SG warnings	27.9%	REF	69.6%	REF	86.1%	REF
1	Addictive	22.8%	0.61 (0.47, 0.78) ^{a, b}	70.0%	1.11 (0.89, 1.39)	86.4%	1.07 (0.78, 1.45)
2	Harm children	40.7%	1.37 (1.11, 1.69) ^{a, b}	83.3%	2.38 (1.82, 3.10) ^{a, b}	83.1%	0.94 (0.70, 1.27)
3	Kill you	34.2%	1.04 (0.83, 1.29)	73.9%	1.70 (1.34, 2.17) ^{a, b}	85.5%	0.99 (0.73, 1.34)
4	Fatal lung disease in nonsmokers	41.9%	1.55 (1.26, 1.91) ^{a, b}	77.3%	1.94 (1.52, 2.49) ^{a, b}	77.5%	0.70 (0.53, 0.92) ^{a, b}
5	Quit now	27.8%	0.95 (0.76, 1.19)	69.5%	1.18 (0.94, 1.47)	87.9%	1.01 (0.75, 1.37)
6	Head and neck cancer	80.9%	8.09 (6.44, 10.16) ^{a, b}	84.5%	2.70 (2.05, 3.55) ^{a, b}	71.6%	0.53 (0.41, 0.68) ^{a, b}
7	Bladder cancer	87.2%	14.63 (11.19, 19.14) ^{a, b}	80.0%	2.14 (1.66, 2.77) ^{a, b}	66.0%	0.43 (0.33, 0.55) ^{a, b}
8	Stunt fetal growth	40.0%	1.73 (1.40, 2.12) ^{a, b}	78.9%	2.00 (1.55, 2.57) ^{a, b}	83.9%	0.93 (0.69, 1.25)
9	Clogged arteries	52.1%	2.64 (2.15, 3.23) ^{a, b}	80.2%	2.05 (1.59, 2.63) ^{a, b}	85.2%	1.14 (0.83, 1.57)

(continued)

Table 3-3. Logistic Regressions of New Information, Thinking about Risks, and Perceived Factualness Comparing GHW with SG Warnings (continued)

Condition	Warning	New Information		Thinking About Risks		Perceived Factualness	
		%	OR (95% CI)	%	OR (95% CI)	%	OR (95% CI)
10	COPD 1 ^c	33.1%	1.48 (1.20, 1.83) ^{a, b}	80.5%	2.25 (1.73, 2.91) ^{a, b}	85.4%	1.23 (0.89, 1.70)
11	COPD 2 ^d	35.7%	1.48 (1.20, 1.83) ^{a, b}	79.2%	2.13 (1.64, 2.75) ^{a, b}	83.8%	1.26 (0.91, 1.73)
12	Erectile dysfunction	78.8%	7.65 (6.10, 9.60) ^{a, b}	77.2%	1.56 (1.23, 1.98) ^{a, b}	72.4%	0.53 (0.41, 0.69) ^{a, b}
13	Amputation	74.7%	7.26 (5.79, 9.11) ^{a, b}	87.5%	3.52 (2.60, 4.75) ^{a, b}	76.7%	0.66 (0.50, 0.86) ^{a, b}
14	Diabetes	87.2%	10.64 (8.34, 13.58) ^{a, b}	76.6%	2.11 (1.63, 2.72) ^{a, b}	64.0%	0.44 (0.34, 0.56) ^{a, b}
15	Macular degeneration	82.6%	11.81 (9.17, 15.21) ^{a, b}	81.4%	2.64 (2.01, 3.46) ^{a, b}	73.7%	0.59 (0.45, 0.77) ^{a, b}
16	Cataracts	88.7%	14.45 (11.08, 18.86) ^{a, b}	75.6%	1.71 (1.34, 2.17) ^{a, b}	65.5%	0.38 (0.30, 0.49) ^{a, b}

^a Significant at $p < .05$ in unadjusted analyses.

^b Significant after adjustment for multiple comparisons.

^c Image of diseased lungs.

^d Image of man with oxygen.

3.2.2 Thinking about Risks

As shown in Table 3-3, 69.6% of control condition participants and a range of 69.5% (Quit now) to 87.5% (Amputation) treatment condition participants thought about the health risks of smoking in response to the warning. GHW were significantly more likely to cause participants to think about the health risks of smoking in 14 conditions relative to the control condition: Harm children, Kill you, Fatal lung disease in nonsmokers, Head and neck cancer, Bladder cancer, Stunt fetal growth, Clogged arteries, COPD 1, COPD 2, Erectile dysfunction, Amputation, Diabetes, Macular degeneration, and Cataracts. These comparisons were significant before and after adjustment.

3.2.3 Perceived Factualness

As shown in Table 3-3, the majority of participants considered the label they viewed to be factual. Specifically, 86.1% of control condition participants and a range of 64.0% (Diabetes) to 87.9% (Quit now) of treatment condition participants believed the label they saw was factual. Participants were less likely to consider the GHW as factual than the control in 8 conditions: Fatal lung disease in nonsmokers, Head and neck cancer, Bladder cancer, Erectile dysfunction, Amputation, Diabetes, Macular degeneration, and Cataracts. These comparisons were significant before and after adjustment for multiple comparisons.

In the ordinal analyses (Tables 3-6 and 3-8), there is a DID score for each level of the outcome variable (strongly disagree, disagree, etc.), and an interaction term for the ordinal model. Significant positive interaction terms, such as 0.50 for Fatal lung disease in nonsmokers in Table 3-6, indicate there was a greater pre-post change in agreement with health beliefs for participants in the treatment condition than in the control condition in a direction that “favors” the GHW (i.e., toward greater agreement with the health belief in treatment over control). Significant negative interaction terms, such as -1.03 for Addictive in Table 3-6, indicate lower agreement with the health belief in treatment over control.

Because of attrition, the sample sizes for Sessions 2 and 3 differ; thus, we present results first for differences between Session 2 and 1 and then differences between Session 3 and Session 1.

3.3.1 Differences in Health Beliefs Between Session 2 and Session 1

As seen in Table 3-5, DID health belief scores were positive and significant before and after adjustment for 9 of the 10 warnings with scores on a linear scale: Head and neck cancer, Bladder cancer, Clogged arteries, COPD 1, Erectile dysfunction, Amputation, Diabetes, Macular degeneration, and Cataracts. That is, there was a net positive increase in agreement with health beliefs for those GHW after accounting for changes in health beliefs in the control condition.

Table 3-5. Linear Regression of Difference in Difference Mean Health Belief Scores between Session 2 and Session 1

Condition	Mean (SD) Health Belief Score		DID (95% CI)
	Session 1	Session 2	
Head and neck cancer	3.40 (3.04)	3.97 (2.94)	0.50 (0.37, 0.63) ^{a, b}
Control	3.35 (1.48)	3.42 (1.46)	
Bladder cancer	3.27 (3.14)	3.96 (2.98)	0.60 (0.47, 0.74) ^{a, b}
Control	3.25 (1.47)	3.34 (1.50)	
Clogged arteries	4.00 (2.67)	4.20 (2.52)	0.18 (0.07, 0.29) ^{a, b}
Control	3.94 (1.32)	3.96 (1.29)	

(continued)

Table 3-5. Linear Regression of Difference in Difference Mean Health Belief Scores between Session 2 and Session 1 (continued)

Condition	Mean (SD) Health Belief Score		DID (95% CI)
	Session 1	Session 2	
COPD 1 ^c	4.35 (2.27)	4.49 (2.17)	0.12 (0.04, 0.21) ^{a, b}
Control	4.36 (1.08)	4.38 (1.09)	
COPD 2 ^d	4.44 (2.36)	4.47 (2.52)	0.01 (-0.09, 0.11)
Control	4.37 (1.08)	4.38 (1.08)	
Erectile dysfunction	3.67 (2.74)	4.16 (2.76)	0.41 (0.28, 0.53) ^{a, b}
Control	3.59 (1.31)	3.67 (1.30)	
Amputation	3.55 (2.96)	4.20 (2.54)	0.56 (0.43, 0.69) ^{a, b}
Control	3.52 (1.46)	3.61 (1.43)	
Diabetes	2.97 (3.12)	3.89 (3.22)	0.74 (0.59, 0.89) ^{a, b}
Control	2.90 (1.52)	3.07 (1.51)	
Macular degeneration	3.32 (2.61)	4.01 (2.89)	0.58 (0.46, 0.70) ^{a, b}
Control	3.21 (1.39)	3.32 (1.43)	
Cataracts	3.14 (3.01)	3.92 (3.06)	0.66 (0.52, 0.80) ^{a, b}
Control	3.09 (1.45)	3.22 (1.45)	

^a Significant at p<.05 in unadjusted analyses.

^b Significant after adjustment for multiple comparisons.

^c Image of diseased lungs. ^dImage of man with oxygen.

Note: Control = average of the 4 Surgeon General warnings for the relevant health belief. Regressions control for age group and smoking status. DID = difference in difference. GHW = graphic health warning. SD = standard deviation. DID scores represent: (Session 2 mean for GHW – Session 1 mean for GHW) - (Session 2 mean for Control – Session 1 mean for Control).

As seen in Table 3-6, DID health beliefs were significantly different before and after adjustment for 3 of the 6 GHW compared with the control condition using ordinal regression: Addictive, Fatal lung disease in nonsmokers, and Stunt fetal growth. That is, the GHW was associated with a greater pre-post change in level of the ordinal dependent variable than the control group (an average of the 4 SG warnings). For the Addictive warning, this difference was in the negative direction (i.e., toward lower levels of agreement with the health belief in the GHW condition compared with the control condition). For the Fatal lung disease in smokers and Stunt fetal growth warnings, the differences were in the positive direction (i.e., toward higher levels of agreement with the health belief in the GHW condition compared with the control condition).

Table 3-6. Ordinal Regression of Difference in Difference Health Belief Scores between Session 2 and Session 1

Warning	Ordinal Response Options for Corresponding Health Belief Item	Condition	Proportion Selecting Response		DID (95% CI)	Interaction Term (95% CI)
			Session 1	Session 2		
Addictive	1 "Strongly disagree"	GHW	0.01	0.01	0.01	-1.03 ^{a, b} (-1.64, -0.41)
		Control	0.02	0.01	(0.00, 0.01)	
	2 "Disagree"	GHW	0.00	0.01	0.00	
		Control	0.01	0.01	(0.00, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.02	0.03	0.02	
		Control	0.04	0.03	(0.01, 0.03)	
	4 "Agree"	GHW	0.13	0.19	0.06	
		Control	0.19	0.19	(0.03, 0.10)	
	5 "Strongly agree"	GHW	0.84	0.76	-0.09	
		Control	0.75	0.76	(-0.15, -0.04)	
Harm children	1 "Strongly disagree"	GHW	0.01	0.01	0.00	0.32 (-0.15, 0.79)
		Control	0.01	0.01	(0.00, 0.00)	
	2 "Disagree"	GHW	0.02	0.02	0.00	
		Control	0.02	0.02	(-0.01, 0.00)	
	3 "Neither agree nor disagree"	GHW	0.07	0.06	-0.01	
		Control	0.07	0.08	(-0.03, 0.00)	
	4 "Agree"	GHW	0.31	0.29	-0.02	
		Control	0.31	0.32	(-0.05, 0.01)	
	5 "Strongly agree"	GHW	0.60	0.62	0.04	
		Control	0.59	0.58	(-0.02, 0.09)	

(continued)

Table 3-6. Ordinal Regression of Difference in Difference Health Belief Scores between Session 2 and Session 1 (continued)

Warning	Ordinal Response Options for Corresponding Health Belief Item	Condition	Proportion Selecting Response		DID (95% CI)	Interaction Term (95% CI)
			Session 1	Session 2		
Kill you	1 "Strongly disagree"	GHW	0.01	0.01	0.00	0.09 (-0.42, 0.61)
		Control	0.01	0.01	(0.00, 0.00)	
	2 "Disagree"	GHW	0.01	0.01	0.00	
		Control	0.01	0.01	(0.00, 0.00)	
	3 "Neither agree nor disagree"	GHW	0.07	0.06	0.00	
		Control	0.07	0.07	(-0.02, 0.01)	
	4 "Agree"	GHW	0.29	0.28	-0.01	
		Control	0.30	0.29	(-0.04, 0.03)	
Fatal lung disease in nonsmokers	1 "Strongly disagree"	GHW	0.03	0.02	-0.01	0.50 ^{a, b} (0.08, 0.92)
		Control	0.03	0.02	(-0.02, 0.00)	
	2 "Disagree"	GHW	0.09	0.06	-0.02	
		Control	0.09	0.08	(-0.04, 0.00)	
	3 "Neither agree nor disagree"	GHW	0.24	0.19	-0.03	
		Control	0.23	0.22	(-0.06, -0.01)	
	4 "Agree"	GHW	0.37	0.38	0.00	
		Control	0.38	0.38	(-0.01, 0.01)	
Quit now	1 "Strongly disagree"	GHW	0.02	0.01	0.00	0.29 (-0.13, 0.72)
		Control	0.02	0.02	(-0.01, 0.00)	
	2 "Disagree"	GHW	0.02	0.02	0.00	
		Control	0.02	0.02	(-0.01, 0.00)	
	3 "Neither agree nor disagree"	GHW	0.08	0.06	-0.01	
		Control	0.08	0.07	(-0.03, 0.01)	
	4 "Agree"	GHW	0.38	0.34	-0.02	
		Control	0.37	0.35	(-0.06, 0.01)	
	5 "Strongly agree"	GHW	0.49	0.57	0.04	
		Control	0.50	0.55	(-0.02, 0.11)	

(continued)

Table 3-6. Ordinal Regression of Difference in Difference Health Belief Scores between Session 2 and Session 1 (continued)

Warning	Ordinal Response Options for Corresponding Health Belief Item	Condition	Proportion Selecting Response		DID (95% CI)	Interaction Term (95% CI)
			Session 1	Session 2		
Stunt fetal growth	1 "Strongly disagree"	GHW	0.01	0.01	-0.01	1.02 ^{a, b} (0.54, 1.49)
		Control	0.01	0.01	(-0.01, 0.00)	
	2 "Disagree"	GHW	0.02	0.01	-0.01	
		Control	0.02	0.02	(-0.02, -0.01)	
	3 "Neither agree nor disagree"	GHW	0.15	0.10	-0.06	
		Control	0.15	0.16	(-0.08, -0.03)	
	4 "Agree"	GHW	0.37	0.32	-0.05	
		Control	0.37	0.37	(-0.08, -0.03)	
	5 "Strongly agree"	GHW	0.44	0.56	0.13	
		Control	0.44	0.43	(0.07, 0.19)	

^a Significant at $p < .05$ in unadjusted analyses. ^b Significant after adjustment for multiple comparisons.

Note: Control = average of the 4 Surgeon General warnings for the relevant health belief. Regressions control for age group and smoking status. DID = difference in difference. GHW = graphic health warning. DID scores represent: (Session 2 proportion for GHW - Session 1 proportion for GHW) - (Session 2 proportion for Control - Session 1 proportion for Control).

3.3.2 Differences in Health Beliefs Between Session 3 and Session 1

As seen in Table 3-7, DID health belief scores were positive and significant before and after adjustment for 6 of the 10 warnings with scores on a linear scale: Head and neck cancer, Bladder cancer, Amputation, Diabetes, Macular degeneration, and Cataracts. There was a net positive increase in agreement with health beliefs for those GHW after accounting for changes in health beliefs in the control condition.

Table 3-7. Linear Regression of Difference in Difference Mean Health Belief Scores between Session 3 and Session 1

Condition	Mean (SD) Health Belief Score		DID (95% CI)
	Session 1	Session 3	
Head and neck cancer	3.22 (2.96)	3.64 (2.85)	0.25 (0.11, 0.40) ^{a, b}
Control	3.36 (1.50)	3.52 (1.45)	

(continued)

Table 3-7. Linear Regression of Difference in Difference Mean Health Belief Scores between Session 3 and Session 1 (continued)

Condition	Mean (SD) Health Belief Score		DID (95% CI)
	Session 1	Session 3	
Bladder cancer	3.25 (3.04)	3.66 (2.87)	0.36 (0.19, 0.52) ^{a, b}
Control	3.26 (1.42)	3.31 (1.49)	
Clogged arteries	3.94 (2.44)	3.99 (2.42)	-0.01 (-0.13, 0.12)
Control	3.95 (1.29)	4.01 (1.31)	
COPD 1^c	4.35 (2.33)	4.34 (2.26)	-0.08 (-0.19, 0.04)
Control	4.38 (1.05)	4.45 (1.05)	
COPD 2^d	4.46 (2.18)	4.43 (2.03)	-0.10 (-0.20, 0.00)
Control	4.39 (1.05)	4.46 (1.05)	
Erectile dysfunction	3.65 (2.81)	3.87 (2.48)	0.10 (-0.05, 0.24)
Control	3.61 (1.31)	3.74 (1.35)	
Amputation	3.46 (2.80)	3.97 (2.56)	0.37 (0.23, 0.51) ^{a, b}
Control	3.55 (1.44)	3.69 (1.46)	
Diabetes	2.90 (2.86)	3.35 (3.04)	0.25 (0.08, 0.42) ^{a, b}
Control	2.89 (1.50)	3.09 (1.53)	
Macular degeneration	3.22 (2.54)	3.60 (2.75)	0.26 (0.13, 0.40) ^{a, b}
Control	3.22 (1.40)	3.35 (1.42)	
Cataracts	3.02 (2.69)	3.55 (2.62)	0.33 (0.18, 0.49) ^{a, b}
Control	3.08 (1.44)	3.27 (1.51)	

^a Significant at p<.05 in unadjusted analyses. ^bSignificant after adjustment for multiple comparisons.
^cImage of diseased lungs. ^dImage of man with oxygen.

Note: Control = average of the 4 Surgeon General warnings for the relevant health belief. Regressions control for age group and smoking status. DID = difference in difference. GHW = graphic health warning. SD = standard deviation. DID scores represent: (Session 3 mean for GHW – Session 1 mean for GHW) - (Session 3 mean for Control – Session 1 mean for Control).

As seen in the ordinal regression results in Table 3-8, DID health beliefs were significantly different before and after adjustment for multiple comparisons for 1 treatment condition (Fatal lung disease in nonsmokers) compared with the control condition. This difference was in the positive direction (i.e., toward higher levels of agreement with the health belief in the GHW condition compared with the control condition).

Table 3-8. Ordinal Regression of Difference in Difference Health Belief Scores between Session 3 and Session 1

Warning	Ordinal Response Options for Corresponding Health Belief Item	Condition	Proportion Selecting Response		DID (95% CI)	Interaction term (95% CI)
			Session 1	Session 3		
Addictive	1 "Strongly disagree"	GHW	0.02	0.02	0.01	-0.39 (-1.08, 0.30)
		Control	0.02	0.02	(0.00, 0.02)	
	2 "Disagree"	GHW	0.01	0.01	0.00	
		Control	0.01	0.01	(0.00, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.03	0.03	0.01	
		Control	0.03	0.03	(-0.01, 0.02)	
	4 "Agree"	GHW	0.15	0.16	0.03	
		Control	0.17	0.16	(-0.02, 0.07)	
Harm children	1 "Strongly disagree"	GHW	0.01	0.01	0.00	-0.21 (-0.79, 0.37)
		Control	0.01	0.01	(0.00, 0.01)	
	2 "Disagree"	GHW	0.02	0.02	0.00	
		Control	0.02	0.02	(0.00, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.08	0.08	0.01	
		Control	0.08	0.07	(-0.01, 0.03)	
	4 "Agree"	GHW	0.32	0.32	0.01	
		Control	0.32	0.31	(-0.02, 0.05)	
Kill you	1 "Strongly disagree"	GHW	0.01	0.01	0.00	-0.26 (-0.82, 0.30)
		Control	0.01	0.01	(0.00, 0.00)	
	2 "Disagree"	GHW	0.01	0.01	0.00	
		Control	0.01	0.01	(0.00, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.07	0.06	0.01	
		Control	0.07	0.06	(-0.01, 0.03)	
	4 "Agree"	GHW	0.29	0.27	0.02	
		Control	0.29	0.26	(-0.02, 0.05)	
	5 "Strongly agree"	GHW	0.62	0.65	-0.03	
		Control	0.62	0.67	(-0.09, 0.03)	

(continued)

Table 3-8. Ordinal Regression of Difference in Difference Health Belief Scores between Session 3 and Session 1 (continued)

Warning	Ordinal Response Options for Corresponding Health Belief Item	Condition	Proportion Selecting Response		DID (95% CI)	Interaction term (95% CI)
			Session 1	Session 3		
Fatal lung disease in nonsmokers	1 "Strongly disagree"	GHW	0.03	0.02	-0.01	0.59 ^{a,b} (0.10, 1.08)
		Control	0.02	0.02	(-0.02, 0.00)	
	2 "Disagree"	GHW	0.10	0.06	-0.02	
		Control	0.08	0.07	(-0.05, 0.00)	
	3 "Neither agree nor disagree"	GHW	0.25	0.19	-0.04	
		Control	0.23	0.21	(-0.07, -0.01)	
	4 "Agree"	GHW	0.38	0.39	0.00	
		Control	0.39	0.39	(-0.01, 0.02)	
Quit now	1 "Strongly disagree"	GHW	0.02	0.01	0.00	-0.06 (-0.54, 0.42)
		Control	0.02	0.01	(-0.01, 0.01)	
	2 "Disagree"	GHW	0.02	0.01	0.00	
		Control	0.02	0.01	(-0.01, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.08	0.06	0.00	
		Control	0.08	0.06	(-0.02, 0.02)	
	4 "Agree"	GHW	0.39	0.34	0.01	
		Control	0.39	0.34	(-0.04, 0.05)	
Stunt fetal growth	1 "Strongly disagree"	GHW	0.01	0.01	0.00	-0.02 (-0.58, 0.54)
		Control	0.01	0.01	(0.00, 0.00)	
	2 "Disagree"	GHW	0.03	0.03	0.00	
		Control	0.03	0.03	(-0.01, 0.01)	
	3 "Neither agree nor disagree"	GHW	0.14	0.13	0.00	
		Control	0.15	0.13	(-0.03, 0.03)	
	4 "Agree"	GHW	0.36	0.36	0.00	
		Control	0.36	0.36	(-0.02, 0.03)	
	5 "Strongly agree"	GHW	0.45	0.47	0.00	
		Control	0.45	0.47	(-0.07, 0.07)	

^a Significant at $p < .05$ in unadjusted analyses. ^b Significant after adjustment for multiple comparisons.

Note: Control = average of the 4 Surgeon General warnings for the relevant health belief. Regressions control for age group and smoking status. DID = difference in difference. GHW = graphic health warning. DID scores represent: (Session 3 proportion for GHW - Session 1 proportion for GHW) - (Session 3 proportion for Control - Session 1 proportion for Control).

4. Summary and Limitations

The primary purpose of this study was to test if GHW increased public understanding of the negative health consequences of cigarette smoking relative to existing SG warnings. Below we describe key findings. A summary of findings across outcomes appears in Table 4-1.

4.1 Summary of Findings

We compared the 16 GHW (two of which shared a text statement but varied in image) with the control (the mean of the 4 SG warnings) on several outcomes: new information, thinking about risks, self-reported learning, perceived factualness, perceived informativeness, perceived understandability, changes in health beliefs, and recall.

Phase 1 results are reactions to warnings after viewing them displayed on mock cigarette packs and advertisements in Session 1. After adjustment for multiple comparisons, respondents were more likely to state they learned new information for 13 of the GHW conditions relative to the control and less likely to state they learned new information for 1 GHW condition relative to the control. Self-reported learning was higher in 13 GHW conditions and lower in 3 GHW conditions relative to the control condition. Out of 16 GHW conditions, 14 were more likely to make participants think about the health risks of smoking than the control, and 8 GHW conditions were less likely to be seen as factual than the control. Perceived informativeness was higher for 13 GHW conditions and lower for 1 GHW condition relative to the control condition. Perceived understandability was higher for 15 GHW conditions.

Phase 2 results represent changes in health beliefs for GHW relative to changes in beliefs for the control condition. Between Session 1 and Session 2, 11 GHW resulted in greater net positive changes in agreement with health beliefs. One GHW resulted in greater net negative changes in agreement health beliefs. From Session 1 to Session 3, 7 GHW resulted in greater net positive changes in agreement with health beliefs.

Phase 3 results are recall of warnings at Session 3. At a median of 17 days after initial exposure, participants in all 16 treatment conditions were more likely to correctly recall which GHW they had seen than participants in the control condition were to recall which SG warning they had seen.

Overall, relative to the average of the SG warnings, many of the GHW were considered to be new information; lead to thinking about risks; result in greater self-reported learning, perceived informativeness, and perceived understandability; and increase agreement with accurate health beliefs over time. The vast majority (nearly three-quarters or more) of participants believed the warning they viewed was a fact, although half of the GHW were seen as less likely to be factual than SG warnings. Participants who viewed SG warnings

were likely viewing something they had seen in real life, particularly so if they were cigarette smokers. In contrast, participants who viewed GHW were inherently viewing something novel. That novelty may produce skepticism. It is possible, though not necessarily the case, that skepticism of the warnings would decline with repeated exposure or if the warnings were viewed in a "real world," rather than hypothetical, context.

Some of the GHW did not perform as well as others for some measures. In particular, the warnings for Addictive, Kill you, and Quit now were often not significantly different from the control condition or were less likely to result in a given outcome (e.g., they produced lower levels of self-reported learning compared with the controls or did not lead to a net positive improvement in accurate health beliefs between sessions). Participants may have muted responses to these warnings because they viewed the content as obvious; it is widely known that smoking is addictive and can kill you and that quitting has positive health effects, even if individuals do not fully understand the extent of these smoking risks and cessation benefits (e.g., Arnett, 2000; Cummings et al., 2004; Mantler, 2013; Murphy-Hoefer et al., 2004; Weinstein, 1998; Weinstein et al., 2004a; Weinstein et al., 2004b).

4.2 Limitations

Some limitations of this study are common to many online studies. For example, the stimuli being tested (in this case, mock cigarette packs and advertisements) were not displayed in a naturalistic fashion but rather on a computer screen. We minimized the lack of realism by showing mock cigarette packs that were the same size as real cigarette packs and enabling participants to rotate the 3-D image of the pack with the warning on it. Two sessions of exposure to stimuli may not be enough to generate changes in some outcomes.

There are also additional, study-specific limitations. Although the universe of respondents included six subgroups (adolescents susceptible to smoking, adolescent current smokers, young adult current smokers, young adult nonsmokers, older adult current smokers, and older adult nonsmokers), we did not have power to look for within-group differences.

In addition, the survey used a convenience sample rather than a probability sample, and the results are not nationally representative. Generating a representative sample of the size necessary for this study would have been cost prohibitive. In addition, an experimental design does not require a nationally representative sample of these subgroups to demonstrate an effect. The panel choice is driven by the large and diverse membership to allow for targeting of adequate numbers of those in the specified tobacco use status groups and to obtain a reasonable degree of demographic diversity in each of the targeted subgroups and the overall sample. Despite efforts to have the study population reflect the demographic makeup of the larger population, the nature of convenience samples still limits the generalizability of the results from this study. These limitations in generalizability do not affect the internal validity, and thus the conclusions, of the study.